(/)

Graduate Course Descriptions

In addition to the course descriptions on this page, you can also view and download syllabi for previously offered courses (/resources/past-courses-outlines). Not all courses are offered every academic session. For sessional offerings, please see our timetable (https://www.statistics.utoronto.ca/curriculum-courses/fall-winter-2021-2022-timetable).
(*) Course requires instructor approval.

JAS1101H - Topics in Astrostatistics

This graduate-level course provides an introduction to the cross-disciplinary field of astrostatistics and is intended for both astronomy and statistics students. We will cover topics in statistics (e.g., hierarchical Bayesian analysis, time series analysis, and cluster analysis) in the context of their applications to astronomical research (e.g., studies of galaxies, the Milky Way, exoplanets, and stellar populations).

These topics will be covered through two main aspects of the course: 1) peer-instruction and collaboration on a term project, and 2) readings, in-class discussion, and exercises related to current astrostats literature. For the term project, the students will develop practical skills by collaborating in cross-disciplinary teams on a research project in astrostatistics using real astronomical data.

Course credit: 0.5 FCE

STA1007H – Statistics for Life and Social Scientists

Consult the instructor for further details.

Prerequisite: Consult the instructor concerning necessary background for this course.

Course credit: 0.5 FCE

STA1008H – Applied Statistics

This course is intended to graduate students in disciplines other than statistics whose studies involve research design and statistical data analysis. Topics include vocabulary of data analysis and principles of research design, significance tests, Type 1 and 2 errors, power and sample size, simple and multiple linear regression, ANOVA, analysis of correlated data (repeated measures), random effects models, introduction to R and computer-intensive methods (permutation, bootstrapping, simulation), and further topics depending on interests of students or instructor.

Prerequisite: One introductory statistics course, or permission from the instructor. Prior programming experience is helpful but not required.

Course credit: 0.5 FCE

STA2005H – Applied Multivariate Analysis

(also offered as undergraduate course STA437H1)

Practical techniques for the analysis of multivariate data; fundamental methods of data reduction with an introduction to underlying distribution theory; basic estimation and hypothesis testing for multivariate means and variances; regression coefficients; principal components and the partial multiple and canonical cor relations; multivariate analysis of variance; classification and the linear discriminant function. The use of R software should be expected.

Course credit: 0.5 FCE

STA2006H – Applied Stochastic Processes

(also offered as undergraduate course STA447H1)

Discrete and continuous time processes with an emphasis on Markov, Gaussian and renewal processes. Martingales and further limit theorems. A variety of applications taken from some of the following areas are discussed in the context of stochastic modeling: Information Theory, Quantum Mechanics, Statistical Analyses of Stochastic Processes, Population Growth Models, Reliability, Queuing Models, Stochastic Calculus, Simulation (Monte Carlo Methods).

Recommended Preparation: knowledge of probability theory calculus and basic real analysis.

Course credit: 0.5 FCE

STA2016H – Spatial Data Analysis

(also offered as undergraduate course STA465H1)

Data acquisition in the environmental, physical, and health sciences are increasingly spatial, and novel in the sense that specialized methods are required for analysis. This course will cover different types of spatial and spatiotemporal data and their analytic methods. Students will learn a variety of advanced techniques for analyzing geostatistical, areal, and point referenced data. Focus will be placed on visualizing spatial data, choosing the correct method for a specific research question, and communicating analytic results clearly and effectively.

Course credit: 0.5 FCE

STA2047H – Stochastic Calculus

A rigorous introduction to stochastic analysis and its applications. Topics include Brownian motion, continuous time martingale, stochastic integration, stochastic differential equations, diffusions, and further topics depending on the interests of the instructor.

Recommended Preparation: knowledge of real analysis and probability theory at the graduate level.

Course credit: 0.5 FCE

STA2051H-Topics in Numerical Methods in Data Science

Techniques for formulating data science models as optimization problems. Algorithms for solving data science problems including gradient-descent based algorithms and randomized algorithms. Emphasis on scalability and efficiency. Convergence analysis of algorithms. Coverage of both convex and nonconvex optimization.

Course credit: 0.5 FCE

STA2052H-Statistics, Ethics, & Law

Modern statistical methods and data analytics are increasingly informing decisions in law, business, medicine, and public life. While the use of statistics to understand social problems is not new, its pervasiveness in society and the scale of available data available opens up a host of new and/or salient moral problems including, for example, fairness, bias, privacy, equality, transparency, accountability, and accessibility.

In this course, we will combine material from law and philosophy together with recent work in statistics and data science in order to gain a better understanding of how to intelligibly reason about these problems, and how to responsibly and creatively apply statistical methods to complex social problems.

The course will be research/project based and the emphasis will be on using statistics to address complex social problems rather than on memorizing abstract ethical principles for handling or processing data.

Recommended Preparation: Graduate students should have an adequate background in probability and statistics (including the use of R). Familiarity with Bayesian approaches and statistical learning/classification would also be helpful.

Course credit: 0.5 FCE

STA2053H – Special Topics in Applied Statistics

The topics will vary from year to year, providing students with the flexibility to examine a diverse range of subjects relevant to applied statistics and data science. This special topics course is repeatable for credit if taken with a different individual topic.

January 2026:

This course covers a range of statistical methods, covering the application and interpretation of models on a range of different datasets. Topics will include Bayesian inference, generalized linear mixed models, generalized additive models involving nonparametric smoothing, model evaluation and selection. We will also cover some core statistical computing techniques. A large focus of the outcomes on this course will also be on reproducible research, dealing with 'real' data, identifying and dealing with data and modeling issues, and model interpretation and communication.

Course credit: 0.5 FCE

STA2080H – Fundamentals of Statistical Genetics*

(also offered as undergraduate course STA480H1)

Statistical genetics is an important data science research area with direct impact on population health, and this course provides an INTRODUCTION to its concepts and fundamentals. We start with an overview of genetic studies to have a general understanding of its goal and study design. We then introduce the basic genetic terminologies necessary for the ensuing discussion of the various statistical methods used for analyzing genetic data. The specific topics include population genetics, principles of inheritance, likelihood for pedigree data, aggregation, heritability and segregation analyses, map and linkage analysis, population-based and family-based association studies and genome-wide association studies. The flow of the content generally follows that of the "The Fundamentals of Modern Statistical Genetics" by Laird and Lange, and additional materials will be provided. Participating students do not need formal training in genetics, but they are expected to have statistical knowledge at the level of STA303 – Methods of Data Analysis or equivalent.

Course credit: 0.5 FCE

STA2101H – Methods of Applied Statistics I

This course will focus on principles and methods of applied statistical science. It is designed for MSc and PhD students in Statistics, and is required for the Applied Paper of the PhD comprehensive exams. The topics covered include: planning of studies, review of linear models, analysis of random and mixed effects models, model building and model selection, theory and methods for generalized linear models, and an introduction to

nonparametric regression. Additional topics will be introduced as needed in the context of case studies in data analysis.

Course credit: 0.5 FCE

STA2102H – Computational Techniques in Statistics

(also offered as undergraduate course STA410H1)

The goal of this course is to give an overview of some of the computational methods that are useful in statistics. The rst part of the course will focus on basic algorithms, such as the Fast Fourier Transform (and related methods) and methods for generating random variables. The second part of the course will focus on numerical methods for linear algebra and optimization (for example, computing least squares estimates and maximum likelihood estimates). Along the way, you will learn some basic theory of numerical analysis (computational complexity, convergence rates of algorithms) and you will encounter some statistical methodology that you may not have seen in other courses.

Recommended Preparation: Background in statistics, computer programming, and linear algebra can be useful for this course.

Course credit: 0.5 FCE

STA2104H – Statistical Methods for Machine Learning and Data Mining

(also offered as undergraduate course STA414H1)

This course will consider topics in statistics that have played a role in the development of techniques for data mining and machine learning. We will cover linear methods for regression and classification, nonparametric regression and classification methods, generalized additive models, aspects of model inference and model selection, model averaging and tree bassed methods.

Course credit: 0.5 FCE

STA2111H – Graduate Probability I

STA 2111H is a course designed for Master's and Ph.D. level students in statistics, mathematics, and other departments, who are interested in a rigorous, mathematical treatment of probability theory using measure theory. Specific topics to be covered include: probability measures, the extension theorem, random variables, distributions, expectations, laws of large numbers, Markov chains.

Students should have a strong undergraduate background in Real Analysis, including calculus, sequences and series, elementary set theory, and epsilon-delta proofs. Some previous exposure to undergraduate-level probability theory is also recommended.

Course credit: 0.5 FCE

STA2112H – Introduction to Statistical Learning I

This course introduces core statistical learning ideas through a computational framework based on (stochastic) gradient descent methods. Students will build on core probability and convergence concepts to learn how common modeling choices translate into practical estimation and prediction procedures. A range of supervised models (linear, generalized linear, and their flexible extensions) will be presented using a (stochastic) gradient descent approach with different losses (squared, logistic, quantile) and regularization. We also introduce basic decision theory to connect predicted probabilities to actions, such as optimal classification thresholds. Model evaluation and model selection methods will be emphasized, as will careful comparison of models using appropriate performance metrics and diagnostic tools.

Course credit: 0.5 FCE

STA2162H – Statistical Inference I

(also offered as undergraduate course STA422H1)

Statistical inference is concerned with using the evidence, available from observed data, to draw inferences about an unknown probability measure. A variety of theoretical approaches have been developed to address this problem and these can lead to quite different inferences. A natural question is then concerned with how one determines and validates appropriate statistical methodology in a given problem. The course considers this larger statistical question. This involves a discussion of topics such as model specification and checking, the likelihood function and likelihood inferences, repeated sampling criteria, loss (utility) functions and optimality, prior specification and checking, Bayesian inferences, principles and axioms, etc. The overall goal of the course is to leave students with an understanding of the different approaches to the theory of statistical inference while developing a critical point-of-view.

Recommended Preparation: Mathematics-based course on the theory of statistics.

Course credit: 0.5 FCE

STA2163H-Online Learning and Sequential Decision Theory

This course presents mathematical foundations for learning, prediction, and decision making. Unlike in traditional statistical learning, however, our focus will be on notions of optimality that do not rely on stochastic modeling assumptions on data. A primary focus will be on learning from data to compete with a class of baselines predictors / strategies, often referred to as experts. A secondary focus will be on the ability to adapt to the presence or absence of statistical patterns, without presuming at the outset that such patterns will arise. Topics include: regret; prediction with expert advice; the role of the loss function in tight bounds; online classification; online linear and convex optimization; regularization; bandit problems / decisions with limited feedback; minimax optimality and adaptivity; relationships with statistical learning.

Recommended Preparation: Mathematical maturity, including real analysis, linear algebra, and probability theory.

Course credit: 0.5 FCE

STA2201H – Methods of Applied Statistics II

The course will focus on generalized linear models (GLM) and related methods, such as generalized additive model involving nonparametric regression, generalized estimating equations (GEE) and generalized linear mixed models (GLMM) for longitudinal data. This course is designed for Master and PhD students in Statistics, and is REQUIRED for the Applied paper of the PhD Comprehensive Exams in Statistics. We deal with a class of statistical models that generalizes classical linear models to include many other models that have been found useful in statistical analysis, especially in biomedical applications. The course is a mixture of theory and applications and includes computer projects featuring R (S+) or/and SAS programming.

Topics: Brief review of likelihood theory, fundamental theory of generalized linear models, iterated weighted least squares, binary data and logistic regression, epidemiological study designs, counts data and log-linear models, models with constant coefficient of variation, quasi-likelihood, generalized additive models involving nonparametric smoothing, generalized estimating equations (GEE) and generalized linear mixed models (GLMM) for longitudinal data.

Course credit: 0.5 FCE

STA2202H – Time Series Analysis

(also offered as undergraduate course STA457H1)

An overview of methods and problems in the analysis of time series data. Topics include: descriptive methods, filtering and adjustment, spectral estimation, bivariate time series models.

The course will cover the following topics:

- Theory of stationary processes, linear processes
- Elements of inference in time domain with applications
- Spectral representation of stationary processes
- Elements of inference in frequency domain with applications
- Theory of prediction (forecasting) with applications > ARMA processes, inference and forecasting
- Non-stationarity and seasonality, ARIMA and SARIMA processes

Further topics, time permitting: multivariate models; GARCH models; state-space models

Course credit: 0.5 FCE

STA2211H – Graduate Probability II

STA 2211H is a follow-up course to STA 2111H, designed for Master's and Ph.D. level students in statistics, mathematics, and other departments, who are interested in a rigorous, mathematical treatment of probability theory using measure theory. Specific topics to be covered include: weak convergence, characteristic functions, central limit theorems, the Radon-Nykodym Theorem, Lebesgue Decomposition, conditional probability and expectation, martingales, and Kolmogorov's Existence Theorem.

Course credit: 0.5 FCE

STA2212H – Statistical Learning II

This course builds on Statistical Learning I to study modern machine learning methods from a statistical perspective. The course covers classical machine learning approaches and unsupervised learning, then develops probabilistic and representation learning frameworks used in contemporary large-scale systems. Gradient-based likelihood optimization for complex models, variational inference, stochastic variational inference, Bayesian neural networks for uncertainty estimation, and representation learning will be covered. Large language models are introduced as autoregressive probabilistic models trained by maximum likelihood. The course emphasizes reproducible computation with robust model evaluation.

Prerequisite: STA2112H

Course credit: 0.5 FCE

STA2311H - Advanced Computational Methods for Statistics I

This course is part one of a 2-course sequence that introduces graduate students to computational methods designed specifically for statistical inference. This course will cover methods for optimization and simulation methods in several contexts. Optimization methods are introduced in order to conduct likelihood-based inference, while simulation techniques are used for studying the performance of a given statistical model and to conduct Bayesian analysis. Covered topics include gradient-based optimization algorithms (Newton method, Fisher scoring), the Expectation-Maximization (EM) algorithm and its variants (ECM, MCEM, etc), basic simulation principles and techniques for model analysis (cross-validation independent replications, etc), Monte Carlo and Markov chain Monte Carlo algorithms (accept-reject, importance sampling Metropolis-Hastings and Gibbs samplers, adaptive MCMC, Approximate Bayesian computation, consensus Monte Carlo, subsampling MCMC, etc). Particular emphasis will be placed on modern developments that address situations in which the Bayesian analysis is conducted when data are massive or the likelihood is intractable. The focus of the course is on correct usage of these methods rather than the detailed study of underlying theoretical arguments.

Course credit: 0.5 FCE

STA2312H - Advanced Computational Methods for Statistics II

The course will discuss the technical side of statistical methods focusing on two key aspects: optimization and implementation. The first part of the course will introduce necessary background for understanding and devising algorithms for modern statistical methodology. It will cover core concepts and tools from convex optimization such as convexity of sets and functions, Lagrange multipliers method, Newton's method, proximal gradient descent, coordinate descent, alternating direction method of multipliers. In addition, it will include the review of key topics in linear algebra such as matrix and vector norms, quadratic forms and positive semidefinite matrices,

matrix calculus (gradient, Hessian and determinant), matrix decompositions (QR, Cholesky, eigen and singular value). The second part of the course will focus on topics from statistical methodology with an emphasis on computational aspects. The covered concepts will include model assessment and selection (bias-variance trade-off, cross-validation and bootstrap), feature selection (penalized generalized linear models, elastic net, group and fused lasso, least angle regression), dimension reduction (principal component analysis, independent component analysis, factor analysis), data compression (k-means, hierarchical, and spectral clustering). The course will involve a significant practical component, which will include labs and coding assignments where students will master their skills in implementing statistical optimization algorithms.

Course credit: 0.5 FCE

STA2453H – Data Science Methods, Collaboration and Communication

This course is designed to provide graduate students with experience in statistical consulting. Students are active participants in research projects brought to the Statistical Consulting Service (SCS) of the Department of Statistics.

Students are not expected to have had any experience as consultants. The purpose of the course is to provide this experience so that graduates will be better able to function in such an environment when they have completed the course. The course also provides students with the opportunity to become familiar with statistical software packages such as The SAS System. There is supervision and assistance to novice consultants.

Content: There is some classroom instruction at the start of the term, an d meetings occasionally are called to discuss special topics and for students to compare experiences. Students serve as apprentice statisticians and work under the guidance of the instructor and the SCS Coordinator on individual projects. Projects are assigned to students as they come in to the SCS. There are periods of inactivity when there are no projects and other times are very busy. The pattern of work is more like that associated with a business or working environment than a traditional course. While some consideration is taken of other academic demands on students, those enrolling must be aware that work on projects may require precedence at times.

Evaluation: Students will be graded on the quality of their work as stati stical consultants. This involves the ability to do work in a timely fashion, the quality of advice provided and the quality of the presentation of advice and written work to clients.

Recommended Preparation: Students should have taken some applied sta tistics courses such as an undergraduate regression course. Also undergraduate courses in applied statistics, sample survey, design of experiments and time series analysis are recommended but these are not required. Also taking some of the other 2000 level applied statistics courses is recommended as this course will serve as an excellent opportunity to put the content of these courses to work.

Course credit: 0.5 FCE

STA2500H - Loss Models: From Data to Decisions

Parametric distributions and transformations, insurance coverage modifications, limits and deductibles, models for claim frequency and severity, models for aggregate claims, stop-loss insurance, risk measures.

Course credit: 0.5 FCE

STA2501H – Advanced Topics in Actuarial Science

Consult the instructor for further details.

Prerequisite: Consult the instructor concerning necessary background for this course

Course credit: 0.5 FCE

STA2502H – Stochastic Methods for Actuarial Science and Finance

(also offered as undergraduate course ACT460H1)

This course is an introduction to the stochastic models used in Finance and Actuarial Science. Students will be exposed to the basics of stochastic calculus, particularly focusing on Brownian motions and simple stochastic differential equations. The role that martingales play in the pricing of derivative instruments will be investigated. Some exotic equity derivative products will be explored together with stochastic models for interest rates.

Recommended Preparation:

- Knowledge of undergraduate probability theory is necessary.
- Knowledge of basic financial modeling (e.g., binomial trees and log-normal distributions) is useful, but not completely necessary.

Course credit: 0.5 FCE

STA2503H – Mathematical Finance*

This course features studies in derivative pricing theory and focuses on financial mathematics and its applications to various derivative products. A working knowledge of probability theory, stochastic calculus (see e.g., STA 2502), knowledge of ordinary and partial differential equations and familiarity with the basic financial instruments is assumed.

The tentative topics covered in this course include, but is not limited to:

- no-arbitrage and the fundamental theorem of asset pricing,
- binomial pricing models;
- continuous time limits;
- the Black-Scholes model;
- the Greeks and hedging;
- European, American, Asian, barrier and other path-dependent options;
- short rate models and interest rate derivatives;
- convertible bonds;
- stochastic volatility and jumps;
- volatility derivatives;
- foreign exchange and commodity derivatives.

More information: Course Website STA 2503. (http://sebastian.statistics.utoronto.ca/courses/sta-2503-math-finance/) (https://web.archive.org/web/20180725085421/http://www.utstat.toronto.edu/sjaimung/courses/mmf1928/main.htm)

Prerequisite: Knowledge of undergraduate probability theory is necessary. Knowledge of basic financial modeling (e.g., binomial trees and log-normal distributions), introductory stochastic calculus and financial products is useful, but not necessary. This course moves at a faster pace, is more advanced and contains a higher workload than STA2502, only students who are well prepared will be allowed to take this course. It is also distinct from STA 2047 which instead focuses on the mathematics of stochastic analysis. This course requires instructor approval prior to enrolment.

Course credit: 0.5 FCE

STA2505H – Advanced Statistical Modelling in Property & Casualty Insurance

(also offered as an undergraduate course ACT466H1)

Limited fluctuation or American credibility, on a full and partial basis. Greatest accuracy or European credibility, predictive distributions and the Bayesian premium, credibility premiums including the Buhlmann and Buhlmann-Straub models, empirical Bayes nonparametric and semi-parametric parameter estimation. Simulation, random numbers, discrete and continuous random variable generation, discrete event simulation, statistical analysis of simulated data and validation techniques.

Recommended Preparation: Consult the instructor concerning the necessary background for this course

Course credit: 0.5 FCE

STA2555H – Information Visualization

In this course we will study techniques and algorithms for creating effective data visualizations based on principles from graphic design, visual art, perceptual psychology, and cognitive science.This course is targeted both towards students interested in using visualization in their own work, as well as students interested in building better visualization tools and systems.

Course credit: 0.5 FCE

STA2600H – Teaching Statistics

This course provides an introduction to a scholarly approach to teaching statistics in higher education. Emphasis is placed on the use of statistics education research, effective communication of fundamental statistical concepts typically encountered in introductory statistics, alignment of learning outcomes, course activities and assessments, recognition of common misconceptions and how to address them, and effective integration of educational and statistical technologies. No prior teaching experience is necessary.

Course credit: 0.5 FCE

STA2700H - Computational Inference and Graphical Models

This is a reading course primarily meant to sequentially follow a modular course offered in the Department. Its purpose is to offer further supervised study of an advanced topic covered for the ambitious student.

Course credit: 0.5 FCE

STA3000Y – Advanced Theory of Statistics

Please note that STA3000Y F & S can only be taken by PhD students in the Department of Statistical Sciences.

This is the Department's core graduate course in statistical theory. It covers the basic principles of statistical inference, their application to a variety of statistical models, and some generalizations to more complex settings.

Prerequisite:

- STA2112H and STA2212H or equivalent. (STA2111H and STA2211H may be co-requisites).
- Some familiarity with measure theory is very useful. The text includes some supplementary material on this.

Course credit: 1.0 FCE

STA3431H – Monte Carlo Methods

This course will explore Monte Carlo computer algorithms, which use randomness to perform difficult high-dimensional computations. Different types of algorithms, theoretical issues, and practical applications will all be considered. Particular emphasis will be placed on Markov chain Monte Carlo (MCMC) methods. The course will involve a combination of methodological investigations, mathematical analysis, and computer programming.

Prerequisite: Knowledge of statistical inference and probability theory at the advanced undergraduate level, and familiarity with basic computer programming techniques.

Course credit: 0.5 FCE

STA4002H - Advanced Special Topics

This graduate course is changes from term-to-term depending on the topic. This section will be updated to reflect the special topics in terms when this course is offered.

Course credit: 0.5 FCE

STA4246H – Advanced Mathematical Finance

This course focuses on advanced theory and modeling of financial derivatives. The topics include, but are not limited to: HJM interest rate models, LFM and LSM market models; foreign exchange options; defaultable bonds; credit default swaps, equity default swaps and collateralized debt obligations; intensity and structural based models; jump processes and stochastic volatility; commodity models. As well, students are required to complete a project, write a report and present a topic of current research interest.

Prerequisite: STA 2503 or equivalent knowledge.

Course credit: 0.5 FCE

STA4273H – Topics Stats Machine Learning: Minimizing Expectations

The problem of minimizing an expected value is ubiquitous in machine learning, from approximate Bayesian inference to acting optimally in a Markov decision process. Progress on this problem may drive advances in methods for generating novel images, unsupervised discovery of object relations, or continuous control. This course will introduce students to various methodological issues at stake in this problem and lead them in a discussion of its modern developments. Introductory topics may include stochastic gradient descent, gradient estimation, policy and value iteration, and variational inference. The class will have a major project component.

Recommended Preparation: This course is designed to guide students in an exploration of the current state of the art, so that ideally, their course projects can make a novel contribution. A previous course in machine learning such as CSC321, CSC411, CSC412, STA414, or ECE521 is strongly recommended. The only hard requirements are linear algebra, multivariate calculus, probability, and basic programming skills.

Course credit: 0.5 FCE

STA4500H – Statistical Dependence: Copula Models and Beyond

The course discusses modern developments in modeling statistical dependence. Emphasis will be placed on copula models, particularly on conditional copula models that can be used in regression settings.

Tentative topics include:

- Random Effects
- Copula Models for Continuous Data
- Dependence measures
- Types of Dependence
- Conditional copulas for Continuous Data
- Copula Models for Discrete/Mixed Data
- Conditional Copula Models for Discrete/Mixed Data

- Vines

Course credit: 0.25 FCE

STA4501H – Functional Data Analysis and Related Topics

Functional data analysis (FDA) has received substantial attention in recent years, with applications arising from various disciplines, such as engineering, public health, finance etc. In general, the FDA approaches focus on nonparametric underlying models that often assume the data are observed from realizations of stochastic processes with smooth trajectories. This course will cover general issues in functional data analysis, such as functional principal component analysis, functional regression models, curve clustering and classification. An introduction to smoothing methods will also be included at the beginning of class to provide a basic view of nonparametric regression (kernel and spline types) and serve as the basis of FDA approaches. The course will involve some computing and data analysis using R or matlab.

Course credit: 0.25 FCE

STA4502H – Topics in Stochastic Processes

This course will focus on convergence rates and other mathematical properties of Markov chains on both discrete and general state spaces. Specific methods to be covered will include coupling, minorization conditions, spectral analysis, and more. Applications will be made to card shuffling and to MCMC algorithms.

Course credit: 0.25 FCE

STA4505H – Applied Stochastic Control: High Frequency and Algorithmic Trading

With the availability of high frequency financial data, new areas of research in stochastic modeling and stochastic control have opened up. This 6 week course will introduce students to the basic concepts, questions and methods that arise in this domain. We will begin with the classical market microstructure models, understand different theories of price formation and price discovery, identify different types of market participants, and then move on to reduced form models. Next, we will investigate some of the typical algorithmic trading strategies employed in industry for different asset classes. Finally, we will develop stochastic optimal control problems for solving optimal liquidation and high frequency market making problems and demonstrate how to solve those problems using the principles of dynamic programming leading to Hamilton-Jacobi-Bellman equations. Students will also have a chance to work with historical limit order book data, develop Monte Carlo simulations and gain a working knowledge of the models and methods.

Tentative topics include:

- Market Microstructure
- Overview of Stochastic Calculus
- Dynamic Programming & HJB -Dynamics of LOB -Optimal Liquidation
- Market Making

- Risk Measures

Course credit: 0.25 FCE

STA4506H – Non-stationary Time Series Analysis

The course will cover modeling, estimation and inference of non-stationary time series. In particular, we will deal with statistical inference of trends, quantile curves, time-varying spectra and functional linear models related to non-stationary time series. With the recent advances in various fields, a systematic account of non-stationary time series analysis is needed.

Course credit: 0.25 FCE

STA4507H – Extreme Value Theory and Applications

Modeling the behaviour of extreme values is important in a variety of disciplines, from finance to environmental science, since catastrophes almost inevitably arise from extreme conditions. This course will cover both theoretical and applied aspects of extreme value modeling. Some of the topics to be covered are: extreme value types, point process methodology, the Hill and other estimators of the tail index, estimating extreme quantiles, multivariate extremes, estimators of tail dependence.

Course credit: 0.25 FCE

STA4508H – Topics in Likelihood Inference

Inference based on the likelihood function has a prominent role in both theoretical and applied statistics. This course will introduce some of the more recent developments in likelihood-based inference, with an emphasis on adaptations developed for models with complex structure or large numbers of nuisance parameters. Special emphasis will be given to applications in biology and medicine throughout the course. Tentative topics to be covered include: review of likelihood inference and asymptotic results; adjustments to profile likelihood; misspecified models — composite likelihood; partially specified models — quasi-likelihood; properties and limitations of penalized likelihood.

Course credit: 0.25 FCE

STA4509H – Insurance Risk Models I

The aim of this course is to provide an introduction to advanced insurance risk theory. This course covers frequent and severity models, aggregate losses and compound distributions, EM algorithm, Model selection and estimation.

Course credit: 0.25 FCE

STA4510H – Insurance Risk Models II

This course aims to discuss the latest research in insurance risk modelling in general insurance. It covers insurance data analysis, probababilty and statistical models for insurance ratemaking and reserving, and their estimation procedures.

Course credit: 0.25 FCE

STA4512H – Logical Foundations of Statistical Inference

The general mathematics and logical foundations for statistical inference: geometric, algebraic and topological symmetries that arise naturally in the solution to the inference problem, including rigorous comparison of the bayesian and frequentist approaches, and the group theoretic considerations of invariance (algebraic and logical symmetry), both on the sample space as well as on the parameter space (and both either implicit or manifest) that must be taken into account in the analysis. Unusual for the development, but fundamental to the inherent logic of such considerations, the finite-finite case is given special attention in respect of both sample space and parameter space.

Course credit: 0.25 FCE

STA4514H – Modelling and Analysis of Spatially Correlated Data

This is an advanced course in models and methods for spatial data, with an emphasis on data which are not normally distributed. The course will cover different types of random spatial processes and how to incorporate them into mixed effects models for normal and non-Normal data, with maximum likelihood and Bayesian inference used for the two types of data respectively. Spatial point processes, where dare are random locations rather than measurements at fixed locations, will be dealt with extensively. Following the course, students will be able to undertake a variety of analyses on spatially dependent data, understand which methods are appropriate for various research questions, and interpret and convey results in the light of the original questions posed.

Course credit: 0.25 FCE

STA4515H – Multiple Hypothesis Testing and its Applications

A central issue in many current large-scale scientific studies is how to assess statistical significance while taking into account the inherent multiple hypothesis testing issue. This graduate course will provide an in-depth understanding of the topic in the context of data science with a focus on statistical `omics'. We start with an insightful revisit of single hypothesis testing, the building block of multiple hypothesis testing. We then study the fundamental elements of multiple hypothesis testing, including the control of family-wise error rate and false discovery rate. We will also touch upon various more advanced topics such as data integration, selective inference and fallacy of p-values. The course will provide both analytical arguments and empirical evidence.

Students are evaluated based on class participation and one final research report on a suggested or self-selected project related to multiple hypothesis testing.

Course credit: 0.25 FCE

STA4516H – Nonstandard Analysis and Applications to Statistics and Probability

Basic concepts in nonstandard analysis, including infinitesimal and infinite numbers, and descriptions of basic concepts like continuity and integration in terms of these notions. Advanced topics, including Loeb measure theory. Applications to stochastic processes and statistics.

Course credit: 0.25 FCE

STA4517H – Foundations & Trends in Causal Interference

This course introduces the research area of causal inference in the intersection of statistics, social science and artificial intelligence. A central theme of this course will be that without a formal theory of causation, intuition alone can be misleading for drawing causal conclusions. Topics include: potential outcomes and counterfactuals, measures of treatment effects, causal graphical models, confounding adjustment, instrumental variables, principal stratification, mediation and interference. Concepts will be illustrated with applications in a wide range of subjects, such as computer science, social science and biomedical data science.

Course credit: 0.25 FCE

STA4518H - Robust Statistical Methods

This course will give an overview of robust statistical methods, that is, methods that are insensitive to outliers or other data contamination. Topics will include theoretical notions such as qualitative robustness and breakdown point, robust estimation of location (minimax variance and bias) and scale parameters, robust estimation in regression and multivariate analysis, and applications (including in computer vision).

Prerequisite:

- STA2112H
- permission of instructor

Course credit: 0.25 FCE

STA4519H - Optimal Transport: Theory & Algorithms

Optimal transport is a vast subject and has deep connections with analysis, probability and geometry. In recent years optimal transport has found widespread applications in data science (a notable example is the Wasserstein GAN). In this course we offer a balanced treatment featuring both the theory and applications of the subject. After laying down the theoretical foundation including the Kantorovich duality, we turn to numerical methods and their applications to data science. Possible topics include entropic regularization, dynamic formulations, gradient flows, statistical divergences and the W-GAN. Our main reference is the recent book Computational Optimal Transport by Gabriel Peyré and Marco Cuturi.

Prerequisite:

- STA2111H – Graduate Probability I
- STA2211H – Graduate Probability II
- (or permission by the instructor)

Course credit: 0.25 FCE

STA4522H – The Measurement of Statistical Evidence

The concept of statistical evidence is central to the field of statistics. In spite of many references to “the evidence” in statistical applications, it is fair to say that there is no definition of this that achieves broad support in the sense of serving as the core of a theory of statistics. The course will examine the various attempts made to measure evidence in the statistical literature and why these are not entirely satisfactory. A proposal to base the theory of statistical inference on a particular measure, the relative belief ratio, is discussed and how this fits into a general theory of statistical reasoning.

More information: Course Website STA4522 (http://www.utstat.utoronto.ca/mikevans/sta4522/STA4522.html).

Course credit: 0.25 FCE

STA4525H – Demographic Methods

This course provides an overview of the core areas of demography (fertility, mortality and migration) and the techniques to model such processes.

The course will cover life table analysis, measures of fertility and nuptiality, mortality and migration models, and statistical methods commonly used in demography, such as Poisson regression, survival analysis, and Bayesian hierarchical models.

The goal of the course is to equip students with a range of demographic techniques to use in their own research.

Course credit: 0.25 FCE

STA4526H - Stochastic Control & Applications in Finance

The course will introduce students to the basic theory of stochastic optimal control. We will cover both the analytic approach, including an introduction to viscosity solution theory, and the probabilistic approach which is based on BSDE and the stochastic maximum principle. Applications to portfolio optimization and contract theory will be discussed. Prerequisite to this course include (measure-theoretic) probability theory and stochastic calculus.

Prerequisite: (measure-theoretic) probability theory and stochastic calculus

Course credit: 0.25 FCE

STA4527H - Random Matrix Theory & Its Applications

Random matrix theory is now a big subject with applications in many disciplines of science, engineering and statistics. This course will cover fundamental concepts, principal and theory in random matrix theory, orienting towards the needs and interests in statistics. Applications to big data analytics and geometric data analysis are provided.

Course credit: 0.25 FCE

STA4528H - Dependence Modelling

This course introduces the theory of modelling dependence in statistical/stochastic models, including copulas and factor models. Typically, data of joint (rare) events are scarce making dependence modelling highly challenging. In financial and insurance risk management, however rare events are prevalent, and misspecification in the dependence structure may greatly impact risk management decisions.
This course provides, additional to copulas and factor models, an overview of financial risk management including risk measures and regulation, such as the Basel accords, with a focus on dependence modelling. It further covers risk assessment and risk management under dependence uncertainty.

Course credit: 0.25 FCE

STA4529H-Applications of Nonstandard Analysis to Statistics and Probability Theory

Nonstandard analysis provides a rigorous foundation for carrying out mathematical analysis with the aid of infinitesimal numbers and other structures that appear in so-called saturated models of the real numbers. This course introduces nonstandard analysis using concepts and examples from statistics and probability. Topics include: extension, transfer, and saturation; infinitesimal and infinite numbers; hyperfinite sets and measures; hyperfinite models of stochastic processes; nonstandard Bayesian decision theory and connections to frequentism. Background in real analysis, probability theory, and statistics recommended. No background will be assumed in mathematical logic.

Prerequisite: Mathematical maturity, including real analysis and probability theory. MSC with instructor approval.

Course credit: 0.25 FCE

STA4530H – Derivatives for Institutional Investing*

This course explores the replication of financial derivatives from the standpoint of investment banks ("sell-side") and the application of derivatives from the standpoint of pension funds, insurers, hedge funds, mutual funds and private equity funds ("buy-side").
The course is structured into three components:

1. The first module analyses how trading and structuring desks at investment banks use vanilla options to create bespoke payouts for institutional investors, corporates and retail investors.

2. The second module examines how the buy-side uses derivatives for:

- Hedging: e.g., protecting traditional balanced portfolios, managing currency risk ;
- Outperforming benchmarks: currency & equity overlay;
- Expressing “macro” views on equity indices, rates, currencies & commodities;
- Expressing “micro” views on sectors & single stocks;
- And addresses why investor preferences give rise to risk premia, and how derivatives can be structured to take advantage of persistent behavioural biases in the market.

3. The third module synthesizes the key learnings from 1. and 2. into case studies

Course credit: 0.25 FCE

Master’s Research Project Course

Supervised Research Project courses, normally taken as half-courses, are offered based on faculty availability. These courses will provide students with a first exposure to research-level topics and thinking. Students will normally be required to write a substantial report about their work, plus perhaps give a brief oral presentation. Projects may be proposed either by faculty or by students.

To enroll in such a course, a student must first obtain permission from the supervising faculty member and from the Associate Chair, Graduate Studies. For further details, please consult the Associate Chair, Graduate Studies.

Webmail (https://webmail.utoronto.ca/) Maps (http://map.utoronto.ca/) A-Z Directory (https://www.utoronto.ca/a-to-z-directory) U of T Contacts (https://www.utoronto.ca/contacts)
U of T (https://www.utoronto.ca/) Follow Us on Twitter (https://twitter.com/UofTStatSci)
Quercus (https://q.utoronto.ca) ACORN (https://acorn.utoronto.ca)

Contact Us

Department of Statistical Sciences
9th Floor, Ontario Power Building
700 University Ave., Toronto, ON M5G 1Z5 (Map)
(https://www.google.ca/maps/place/700+University+Ave.,+Toronto,+ON+M5G+1Z5)
416-978-3452
Email Us (mailto:info.statistics@utoronto.ca)

(https://www.instagram.com/uoftstatsci/) (https://www.linkedin.com/school/uoftstatsci)

Accessibility Policy (/accessibility) Privacy Policy (/privacy)

© 2026 Faculty of Arts & Science, University of Toronto